



BIG DATA ANALYTICS IN HEALTHCARE INDUSTRY

Lata Virendrabhai Upadhyay

Assistant Professor, M. V. M. College of Comm., Mgt & I.T., Saurashtra University

ABSTRACT

On a daily basis, we refer to the present era as the 'modern age', which can be named as the era of big data in the field of information technology field. In today's world, our daily life is moving fast and never thirsting. The fields of science, science, engineering and technology produce data at a fatal rate, leading to data from Exhibit (s) every day. Big data helps us explore and rediscover many areas where education, health and law are not limited. The main purpose of this paper is to provide in-depth analysis of healthcare in the field of healthcare using big data and analytics. The main purpose is to emphasize the use of big data that helps look back in history, but this is the time to focus on analytics to improve medicine and services. Although many big data implementations are intended for home development, this proposed implementation is intended to propose a wider range using Hadoop, which is intended only as the tip of the iceberg. The focus of this paper is not just limited to the correction and analysis of data; It also focuses on strengths and drawbacks compared to traditional techniques available

KEYWORDS: Big data, Healthcare, Information Technology.

1. INTRODUCTION:

Data is a powerful tool that comes in many forms. Big data does not have a global definition, when it is discussed differently. The term Big Data is referred to in various fields to describe the deadly development of data flows that are too large to process using the traditional database and software software techniques available. Although big data is often considered scary, it is a blast in the field of information. It helps to perform various analyzes, which can impact economic development, create opportunities, improve efficiency on other organizations.

2. WHAT IS BIG DATA ANALYTICS:

Data is the oil of the 21st century, and examination is the ignition engine. — Peter Sondergaard, Gartner Research. A huge development is recognized when it can have a constructive outcome. The information gathered in different vaults by different associations, the information created by the people can have any kind of effect just on the off chance that we can break down and use the information appropriately. At the end of the day, without legitimate examination, information will be only an asset however not a used asset. Moreover, here the term enormous information does discuss the volume of information as well as the intensity of the information. The informational indexes are huge and perplexing, moving the present methods to break down and catch the results. So as to survive and settle on choices in the quickly developing business area, the huge information investigation filters through the information to reveal the concealed example, acknowledge obscure relationships, understanding the market patterns, client inclinations and other valuable business data.

3. BIG DATA IN HEALTHCARE:

Characterizing the volume of the information, the sort of information, and the element confinements are wide. Consolidating the total medicinal services information is huge, which alludes as —big data yet not as gigantic, as we imagine that the association can't deal with its information. The vast majority of the healthcare suppliers have not confronted any more difficult circumstance to deal with the information for them; be that as it may, it is in every case great to anticipate the innovation enhancements and execution, which can support them. As indicated by the McKinsey Global Institute, better focusing of precaution healthcare messages to the correct populace at the opportune time could spare \$70-100 billion. Subsequently, Hadoop information preparing is the outstanding amongst other decision to go with at the present patterns. The computational capacities of Hadoop preparing will probably invigorate the numerical techniques accessible right now, medicinal research ways to deal with increment the result quality. Albeit enormous information probably won't be their case, yet the revelation procedure to discover new strategies to break down the information they have, increment the exactness of the exploratory outcomes, give different instruments to locate the nature of information is consistently at the most noteworthy need. Subsequently, Hadoop information handling is the outstanding amongst other decision to go with the present patterns. The computational abilities of Hadoop preparing will most likely animate the numerical techniques accessible as of now, and therapeutic research ways to deal with increment the result quality. —Most of the information frameworks are for charging, and they aren't utilized to improve the nature of care, clarifies Jason Jones, official chief for clinical knowledge and choice help at Kaiser Permanente, a human services supplier and not-revenue driven wellbeing plan that serves roughly 9.1 million individuals in 8 states and the District of Columbia. The developing conventional human services frameworks typically spare and oversee EMR (Electronic Medical

Record), PHR, Laboratory Information System (LIS), biomedical information, biometrics information, and genomic information which can be the significant sources to create the results. These different information sources help to process and examine the information with different qualities. Handling such gigantic informational indexes utilizing the Hadoop innovation will encourage us not exclusively to process immediately contrasted with the conventional database arrangements which are being utilized right now, yet additionally gives an additional edge to investigate the information distinctively.

4. OBJECTIVES:

The objective of my project is to propose a feasible computing solution using the big data and analytics. It aims to foster the research, availability and accessibility in the field of health care. This project also provides measurable benefits providing the ground realities to improve the field of health care. It targets to benefit the society with advanced computation techniques to analyze and provide patient-centric health care. The proposed objectives with detailed description are as below.

4.1 Clinical Decision Support:

The Clinical Decision Support (CDS) aims to increase the quality of health care services enhancing the outcomes. The primary focus of the system is to provide the right information to the right people, proper customized health care management process but not limited to clinical guidelines, documentations, and diagnosis. This system enables doctors, pharma, patients and other individuals to know the information in any specific medical related information.

4.2 Disease Management:

This system enables to analyze various diseases, its evolution using the laboratory tests. The analytics help to improve the accuracy to find out the information enabling to improve the outcomes such that medical solutions can be determined. This requires the data support from various organizations, medical repositories and the individuals.

4.3 Patient Matching:

With the help of prescriptive analytics using the big data solutions, patient-centric medical approach is developed. It aims to explore the previous disease management systems, the approach to cure the patient, symptoms while fighting the disease etc. By exploring and analysis, the outcomes could be clear enough to treat a patient based on the symptoms instead of a generic disease-based management.

4.4 Lifestyle Analytics:

The proposed system helps to provide the health care solutions based on various methods but also life style of the individuals. In order to prevent the medical accidents and increase the accuracy towards the disease detection, lifestyle of the individual plays an important role. This system provides the picture in a wide range of various lifestyle activities done by the individuals. This helps to study the impacts and the causes.

5. DATA MATCHING:

In order to implement the system without any errors, as accurate as possible and achieve greater results; it is required to deal with multiple repositories, and medical organizations.

6. DATA SECURITY:

Throughout the process, Privacy will be the primary concern. Since, medical records contain various personal information; data access is limited to the authorized users. Besides, data is stored in a highly secured server including regulations such as Health Insurance Portability and Accountability Act of 1996 (HIPAA) and the Data Protection Act and encryption techniques.

7. DATA PYRAMID:

The pyramid depicts the architecture and explains about the Data Management. It clearly gives the picture of the data flow starting from the raw data along with its types through the Hadoop eco-system and the analytic engines to achieve the final goal of the system.

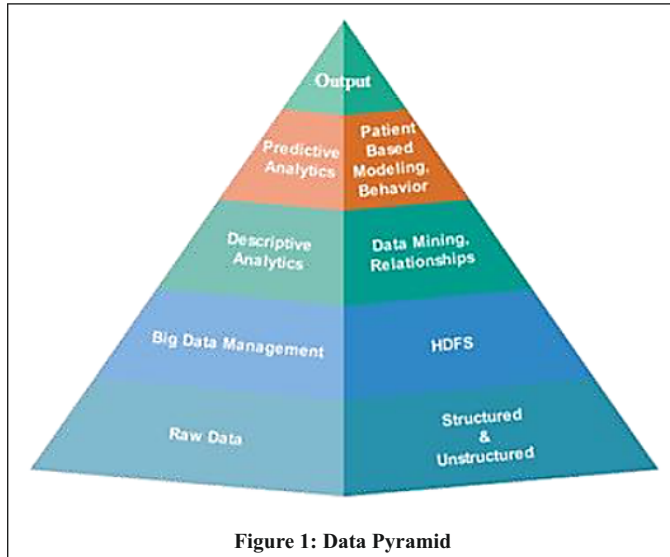


Figure 1: Data Pyramid

8. DATA LIFECYCLE:

The lifecycle of a system defines the process that has ability to impact on the efficiency providing problem solutions.

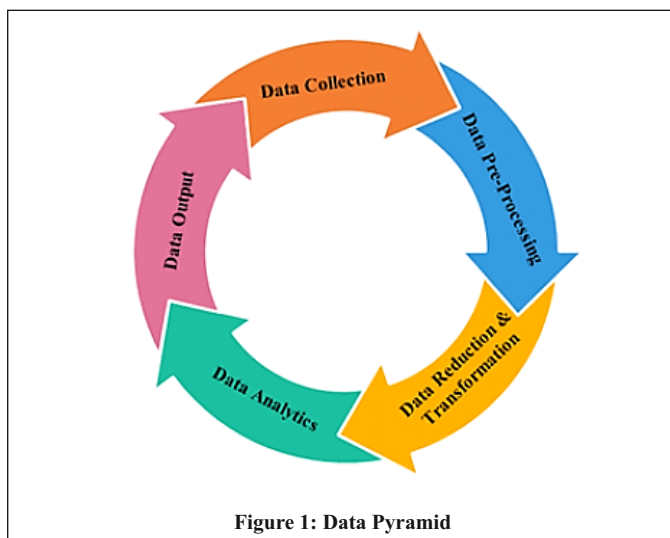


Figure 1: Data Lifecycle

8.1 Data Collection:

The very first stage of the process is collecting the data from various repositories, sources and storing it in Hadoop Distributed File System (HDFS). Data can be both structured and unstructured like clinical analysis, pharmaceutical purchases, patient history, reports, medical emergencies, health index, and social media content, wearable devices and so on. The impact on the quality can be measured with the collection process.

8.2 Data Pre-processing:

As we collect the data from different sources and store it into a common storage, this stage helps to cleanse the data, sort accordingly for further analysis. All the missing values, columns or data will be discarded. Preparing the data plays an important role in moving the process ahead as raw data cannot be processed and might not help to achieve any results. In this process, any junk data will be discarded.

8.3 Data Reduction & Transformation:

Thus far we have been staging & preparing the incoming data. This stage helps to process the data from the pipeline by reducing least prominent data (or) columns. Data transformation aims to transform the data applying various mathematical &

compression algorithms. Transforming the data will not lose any data rather helps to sort out to process further easily. Various analytics and analysis functions can be implemented on the transformed data.

8.4 Data Analytics:

Data Analytics plays the major role in analyzing the data by drawing conclusions. It interprets the data for research, science and business decisions. Using these outcomes, business models can be improved, unknown correlations can be made.

8.5 Data Output:

The statistical, analytical information is presented based on the inputs given. The generated reports are interpreted appropriately in various formats for the user to view, monitor or print. The output data is the meaningful information that guides and helps to achieve the goal. In order to store the output data for any further use, data storage is required. Similarly, it is important to maintain security, integrity and data access control specified throughout the lifecycle.

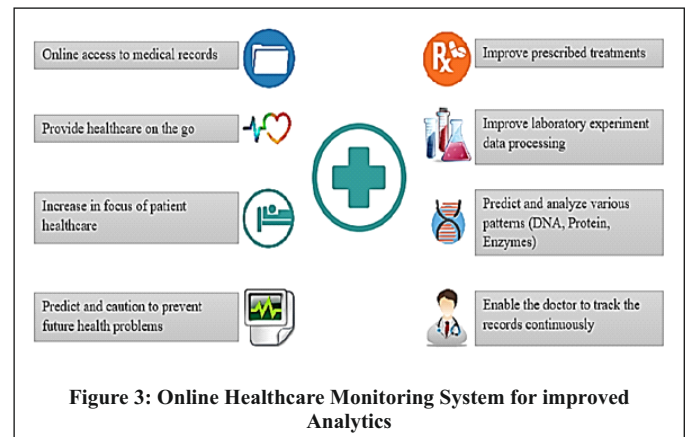


Figure 3: Online Healthcare Monitoring System for improved Analytics

9. CONCLUSION:

One of the essential budget oriented talk is healthcare and its cost, mainly in the countries having very less economic status, areas with bad health hygiene, increasing births and aged people. In this study, we propose to analyze the healthcare using the big data analytics specifically to any given geographic location and the data available. In addition, moving towards the big data storage and solutions would provide an efficient solution in contrast to the traditional storage solutions. Any further research can easily extend the system to improve the facilities and services.

REFERENCES:

- I. Brian Eastwood. (2013). 6 Big Data Analytics Use Cases for Healthcare IT. [Personnel Survey] Unpublished raw data. Retrieved from <http://www.cio.com/>
- II. Emad A Mohammed, Behrouz H Far, and Christopher Naugler. (2014). Applications of the MapReduce programming framework to clinical big data analysis: current landscape and future trends, Mohammed et al. BioData Mining 2014, 7:22.
- III. Harvard Business. (2014). How Big Data Impacts Healthcare by Harvard Business Review Analytics Services.
- IV. Hongyong Yu, Deshuai Wang. (2012). Research and Implementation of Massive Health Care Data Management and Analysis Based on Hadoop, Fourth International Conference on Computational and Information Sciences.
- V. IBM. (2013). Descriptive, predictive, prescriptive: Transforming asset and facilities management with analytics, Retrieved from document. TIW14162USEN, IBM Corp.
- VI. Lars George. (2014). Getting Started with Big Data Architecture. Cloudera. Retrieved from URL <http://blog.cloudera.com/blog/2014/09/getting-started-with-big-data-architecture/>
- VII. Monya Baker, et al. (2012) Functional Genomics: The changes that count. Nature and Nature Methods Vol. 482, 258-262.
- VIII. Sa-kwang Song et al. (2013). Prescriptive Analytics System for Improving Research Power, IEEE 16th International Conference on Computational Science and Engineering, pp.1144-1145.